

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Sektor industri pariwisata menjadi salah satu sektor yang tumbuh pesat di seluruh dunia tanpa diragukan lagi. Indonesia, yang merupakan anggota kawasan ASEAN (*Association of Southeast Asian Nations*), sering dijuluki sebagai "surga pariwisata" dan memiliki pertumbuhan industri pariwisata yang paling pesat (Sabon dkk. 2018). Perkembangan teknologi dan *platform* digital telah memberikan dampak signifikan pada industri pariwisata, termasuk dalam hal penyebaran berita terkait pariwisata. Informasi tentang pariwisata di Indonesia mudah ditemukan di berbagai media massa, baik cetak maupun digital. Saat ini, masyarakat cenderung mencari informasi pariwisata lewat internet atau media sosial karena lebih mudah, murah, dan cepat dibandingkan dengan media tradisional seperti koran dan majalah (Mursalina dkk. 2019).

Portal berita online merupakan salah bentuk dari media massa digital yang digunakan oleh hampir seluruh kalangan masyarakat. Kemunculan portal berita online ini sejalan dengan perkembangan masyarakat yang semakin dinamis dalam mencari informasi di media massa (Setiowati dkk. 2021). Berdasarkan hasil survey yang dilakukan *Reuters Institute* terhadap sumber berita masyarakat Indonesia didapatkan bahwa media online menjadi sumber informasi yang paling populer bagi

masyarakat Indonesia dengan persentase 89% responden di tanah air mengakses berita melalui portal media online, 64% mengakses berita melalui *platform* media sosial, 58% mengakses berita melalui televisi dan 20% responden mengakses berita melalui media cetak seperti koran dan majalah (Mahdi 2022).

Portal berita online dalam menyajikan beritanya menyediakan ruang untuk berita-berita tentang pariwisata. Banyaknya artikel berita tentang pariwisata di Indonesia yang diterbitkan di portal media online terkadang menimbulkan dampak positif atau negatif bagi masyarakat yang membacanya yang awalnya beropini positif bisa menjadi negatif ataupun sebaliknya dalam skala besar (Rameshbhai dkk. 2019). Dengan banyaknya artikel berita tentang pariwisata di Indonesia peneliti mencoba untuk melakukan klasifikasi sentimen melalui berita-berita yang dipublikasikan tentang pariwisata di Indonesia yang ada di portal media online.

Dalam penelitian ini, klasifikasi sentimen dilakukan pada judul berita, menurut (Dor 2003) sebagian besar orang menilai isi berita secara langsung dengan hanya membaca judul berita secara relatif dari pada membaca keseluruhan berita. Selain itu judul berita dapat mengarahkan pembaca ke ideologi tertentu yang tergambar dalam isi berita dan dianggap mewakili ideologi media yang bersangkutan terhadap apa yang diharapkan (Rizam 2019).

. Berita yang dipublikasikan melalui portal berita online merupakan salah satu bentuk media dalam mengekspresikan pendapatnya tentang beragam kejadian, informasi, fenomena dan fakta mengenai pariwisata yang ada di Indonesia. Dimana berita-berita tersebut bersifat penilaian subjektif atau opini dengan tetap berpedoman pada etika jurnalistik (Juditha 2017). Berita yang dipublikasikan

kemungkinan akan menimbulkan persepsi baru dari masyarakat yang membacanya yang dapat mempengaruhi citra destinasi pariwisata yang di beritakan.

Oleh karena itu, penelitian ini menawarkan solusi dengan membuat sebuah model *classifier* untuk mengklasifikasikan sentimen atau sikap emosional yang terkandung dalam judul berita serta membandingkan metode ekstraksi fitur *Binary Term Presence* dan *Term Frequency*. Dalam penelitian ini, perbandingan antara kedua metode ekstraksi fitur tersebut akan dilakukan untuk mengetahui metode ekstraksi fitur mana yang lebih efektif dalam mengklasifikasikan sentimen judul berita. Apakah menggunakan *Binary Term Presence* yang hanya memperhatikan keberadaan kata dalam judul, atau menggunakan *Term Frequency* yang juga memperhitungkan frekuensi kemunculan kata-kata tersebut.

Sehingga membantu pemerintah maupun *stakeholder* yang terkait dalam memperoleh wawasan dan informasi yang mendukung tentang reaksi masyarakat terhadap topik atau isu pariwisata yang diangkat oleh media. Sehingga dapat digunakan untuk mendukung dalam pengambilan keputusan terkait aspek pariwisata Indonesia karena hal ini berdampak sangat besar bagi citra destinasi pariwisata Indonesia dimata masyarakat (Khasanah, S. R., dkk, ,2022).

Website Detik.com merupakan situs berita utama yang masuk dalam 10 besar *website* yang paling sering dikunjungi oleh masyarakat Indonesia berdasarkan data Alexa maret 2020. Detik.com memiliki beberapa beberapa kanal berita seperti detikNews, detikSport, detikFinance, detikHot, detikTravel, dll. (Rumata 2017) menyatakan bahwa Detik.com adalah portal berita *online* yang masih menjaga netralitas dalam pemberitaannya. Selain itu, (Hutami dkk. 2022) juga menyatakan

bahwa portal berita Detik.com cenderung menjaga stabilitas pemberitaan sehingga tidak menimbulkan spekulasi berlebihan yang belum tentu sesuai fakta.

Pada penelitian ini, peneliti akan membuat sebuah model atau *classifier* mengklasifikasikan sentimen judul berita pariwisata di Indonesia ke dalam kelas sentimen positif, netral dan negatif dengan sumber data dari portal berita online. Dengan metode yang digunakan pada penelitian ini adalah *Support Vector Machine (SVM)*. Peneliti memilih metode SVM dikarenakan SVM merupakan metode yang tepat dalam mencari pemisah antar kelas atau *hyperplane*, dimana SVM akan melakukan klasifikasi dengan membangun di ruang dimensi yang lebih tinggi dan membagi data secara optimal ke dalam kelas-kelas pada *hyperplane* (Adankon dan Cheriet, 2015) dan berdasarkan beberapa penelitian sebelumnya yang telah menunjukkan bahwa metode SVM bekerja jauh lebih baik dalam mengklasifikasikan dokumen teks dibandingkan metode klasifikasi tradisional *machine learning* lainnya seperti *K- Nearest Neighbor (KNN)* dan *Naïve Bayes* (Nanda dkk. 2022).

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, adapun rumusan masalah dalam penelitian ini sebagai berikut:

1. Bagaimana mendapatkan *classifier* untuk mengklasifikasikan judul berita pariwisata Indonesia berdasarkan sentimen ?
2. Bagaimana hasil perbandingan akurasi ekstraksi fitur *Binary Term Presence* dan *Term Frequency* ?

1.3 Tujuan Penelitian

Berdasarkan permasalahan yang telah di paparkan pada rumusan masalah, adapun tujuan dalam penelitian ini sebagai berikut:

2. Menghasilkan *classifier* untuk mengklasifikasikan judul berita pariwisata Indonesia berdasarkan sentimen.
3. Mengetahui hasil perbandingan akurasi ekstraksi fitur *Binary Term Presence* dan *Term Frequency*.

1.4 Batasan Masalah

Agar penyusunan penelitian ini tidak keluar dari pokok permasalahan dan agar tujuan pembahasan semakin terarah maka dilakukan batasan masalah dalam penelitian ini sebagai berikut:

1. Metode yang digunakan dalam pengklasifikasian ini adalah *Support Vector Machine* (SVM) dan tidak membandingkan dengan metode lain.
2. Data yang di gunakan adalah data judul berita yang di kumpulkan dari halaman website Detik.com dengan kata kunci 'pariwisata ,wisata, liburan'
3. Label pada penelitian ini menggunakan 3 label yaitu positif, netral, dan negatif.
4. Penilaian opini positif, netral dan negatif pada data akan dilakukan oleh mahasiswa tingkat akhir yang memiliki kompetensi di bidang bahasa Indonesia.
5. Kernel yang digunakan pada klasifikasi SVM yaitu kernel RBF, Polinomial dan Sigmoid.
6. Data diambil dari bulan Januari – Desember tahun 2022.
7. Visualisasi kata terbanyak digunakan dalam judul berita dengan *WordCloud*.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini dibedakan menjadi dua jenis yaitu manfaat teoritis dan manfaat secara praktis sebagai berikut :

1.5.1 Manfaat Teoritis

Penelitian ini dapat digunakan sebagai referensi atau bahan acuan untuk peneliti lainnya yang memiliki penelitian sejenis dimasa yang akan datang.

1.5.2 Manfaat Secara Praktis

a. Bagi Mahasiswa

Adapun manfaat bagi peneliti dari dilakukannya penelitian ini adalah untuk memenuhi tugas akhir serta sebagai acuan bagi mahasiswa lain yang memiliki penelitian sejenis mengenai pembangunan *classifier* untuk klasifikasi menggunakan metode SVM.

b. Bagi Pemerintah

Pengklasifikasian sentimen data berita tentang pariwisata di Indonesia untuk memudahkan pemerintah atau pihak *stakeholder* yang bersangkutan memperoleh wawasan dan informasi yang mendukung tentang reaksi masyarakat terhadap topik atau isu pariwisata yang diangkat oleh media. Sehingga dapat di gunakan untuk mendukung dalam pengambilan keputusan terkait aspek pariwisata Indonesia.